

Peabody, Daniel (EGLE)

From: Peabody, Daniel (DEQ)
Sent: Monday, December 10, 2018 9:34 PM
To: Keiser, Jeff/MKE; Karen Kirchner (kirchner.karen@epa.gov)
Cc: Saric, James; Kirchner, Scott; Roberts, Keegan; John Kern
Subject: Area 4 Mapping
Attachments: Area 4 PCB Mapping Evaluation 2018_12_10.pdf

Jeff and Karen,

During the call last week regarding the Area 4 mapping, Dr. Kern identified some concerns with the variograms and approaches being utilized by Wood. Over the last few days, Dr Kern prepared a very thorough and thoughtful analysis that shows how we can combine the RI and SRI datasets and identifies some of the issues that are present in the statistical analysis and maps that were completed and presented by Wood.

Thanks,

Daniel Peabody

Environmental Quality Analyst
Site Assessment Site Management Unit
Superfund Section
Remediation and Redevelopment Division
Michigan Department of Environmental Quality
Constitution Hall
525 W Allegan St.,
Lansing, MI 48929
Office: (517) 284-5072
Email: PeabodyD@michigan.gov
Office Hours: T,TR,F 0715-1545, MW 0900-1730

ALLIED PAPER, INC./PORTAGE CREEK/KALAMAZOO RIVER SUPERFUND SITE

AREA 4 MAPPING EVALAUTION

AROCLOR CONCENTRATIONS IN SURFICIAL FORMERLY
IMPOUNDED SEDIMENTS AT TROWBRIDGE IMPOUNDMENT

Draft

December 10, 2018

CONTENTS

2	Overview	4
3	Preview of Findings	5
4	Review of Technical Workgroup Information:.....	6
4.1	Intuitive Appeal of SRI only Ordinary kriging Model.....	7
4.2	Semivariogram Analysis	9
4.2.1	Which Model to Trust	14
4.2.2	Role of the Semivariogram in Kriging.....	17
4.2.3	Practical Implications	17
5	Alternative Mapping	18
5.1	Model Fit (Cross Validation).....	18
5.2	Area Exceeding 11 mg/kg Total Aroclors	19
6	Sensitivity to Aroclor Correction Factors	23
7	Discussion and Conclusions	28
8	References	30

List of Figures

Figure 1. ordinary kriging model for SRI data (top panel) and RI+SRI combined data (bottom panel) reproduced for technical discussion purposes.	8
Figure 2. Example semi-variogram estimates and fitted model defining key parameters.....	10
Figure 3. Sample semivariogram grouping options for estimating a directional semivariogram, reproduced from Deutch and Journel (1998).	11
Figure 4. Empirical semivariogram estimates and models for the SRI data (left panel) and for combined RI and SRI data (right panel), reproduced for discussion purposes from working group presentations. Vertical axis scales differ as presented in original materials.	12
Figure 5. Semivariogram estimates and fitted model for the long flow direction with a range of influence of approximately 125 units and a 14% nugget effect. Inset provided for comparison to GP model presented at working group meetings indicating 50% nugget.	13
Figure 6. Combined RI and SRI sampling locations represented in the longflow (I coordinate) and cross flow (J coordinate) directions.	15
Figure 7. Long- flow (left panel) and cross-flow (right panel) semivariogram models for Aroclors illustrating an approximately 4 to 1 ratio of anisotropy.	16

Figure 8. Natural neighbor interpolation of Aroclors for SRI data only in Area 4 floodplain surface sediments.	21
Figure 9. Natural neighbor interpolation of Aroclors for combined RI and SRI data in Area 4 floodplain surface sediments.	22
Figure 10. Natural neighbor interpolation of total Aroclors for combined RI and SRI data with a 25% correction (Factor of 1.25) applied to the SRI data in Area 4 floodplain surface sediments.	25
Figure 11. Natural neighbor interpolation of total Aroclors for combined RI and SRI data with a 50% correction (factor of 1.5) applied to the SRI data in Area 4 floodplain surface sediments.	26
Figure 12. Natural neighbor interpolation of total Aroclors for combined RI and SRI data with a 100% correction (Factor of 2) applied to the SRI data in Area 4 floodplain surface sediments.	27
Figure 13. Empirical long-flow semivariogram estimates and fitted models for total Aroclors based on RI and SRI data combined. Nugget effect was estimated considering pairs of samples within 0.28 units and 0.44 units irrespective of direction and remaining pairs of points were constrained to be within an angular tolerance of 12 degrees (atoll parameter in GSLIB) and a directional band width (bandh parameter in GSLIB) of 2 units.	29

List of Tables

Table 1. Mean absolute cross validation errors for ordinary kriging model based on SRI data exclusively and RI+SRI data combined.	7
Table 2. Model Fit and Area Estimates for PCB Concentration Exceeding 11 mg/kg Total Aroclors	19
Table 3. Sensitivity of remedial footprint estimates to correction of supplemental RI data.	24

1 OVERVIEW

Accurately determining the area of formerly impounded sediments with total polychlorinated biphenyl (tPCB) concentrations exceeding 11 mg/kg is a key component of Allied Paper, Inc./Portage Creek / Kalamazoo River Superfund Site (Site) Area 4 remedial investigation (RI) and feasibility study (FS). Generally speaking, the remedial footprint size and corresponding remedial cost are approximately proportional to the area exceeding 11 mg/kg tPCB. Investigations of this formerly impounded floodplain have included data collected in 1994, 2000, 2012 and 2014. Surface floodplain soil (formerly impounded sediments) samples were collected from a total of 411 unique locations: 398 collected by consultants working on behalf of Georgia Pacific (GP), and 13 collected by consultants working on behalf of the State of Michigan (State). Of those samples collected by GP consultants, 102 locations were sampled by Arcadis, Inc. as part of the sitewide RI and 296 locations were sampled by Wood in 2014 and referred to as the supplemental remedial investigation (SRI) samples. GP has proposed developing remedial footprints either based solely on the SRI data or weighting the SRI and RI datasets differently. Such a proposal would either remove or artificially bias approximately 25% of the data set collected and relied upon within Area 4 and throughout the site since 1994. At this time, GP has provided no definitive or defensible explanations for why the apparent lower SRI total Aroclors (calculated from Aroclors 1016 through 1260, and referred to as “Aroclors” in this document) data appear to be lower than the RI Aroclor data. Similar apparent inconsistencies between SRI and RI Aroclor concentrations have also been identified at Areas 1, 2 and 3. However, at Areas 1 and 2, there are also Aroclor data collected by USEPA in 2001 that are consistent with the 1994 RI data collected by GP (Kern Statistical Services, Inc., 2001).

The Michigan Department of Environmental Quality’s (MDEQ) contract laboratory recently analyzed Area 1 floodplain soil split samples and certified reference materials (CRM). The contract laboratory’s results suggest that Aroclor totals reported by GP are slightly lower than MDEQ’s split samples, and that the MDEQ Aroclor totals were biased low relative to the CRM (CDM, 2018a). In addition, the MDEQ has also subjected these same split samples to a congener-based analysis. These congener-based analyses indicate that the Aroclor results reported by GP underrepresent actual tPCB concentration. Based on a small number of split samples, the ratio of tPCB to Aroclors reported by GP is estimated to be on the order of 1.25 to 2.0 (CDM, 2018b). Analysis of more split samples are needed to precisely estimate this relationship, so we report ranges of results in the remainder of this document.

As previously stated, GP has not identified any particular reason for the discrepancy between the RI and SRI Aroclors but has proposed basing the FS cost and volume estimates primarily or even exclusively on the SRI data. Their proposal includes removing the RI data from the analysis or restricting its influence in some ad-hoc manner. MDEQ’s position remains that the RI and FS should be based on all data collected over the RI/SRI period of time (1994 through 2014) and that best efforts should be made to determine the accuracy of both data sets through a rigorous evaluation of split samples comparing Aroclors with tPCB congeners by EPA Method 1668 (USEPA, 1999).

In this report we provide three evaluations that we have conducted:

- 1) a review of selected analyses presented by GP supporting the development of the FS based solely on the SRI data,
- 2) a re-evaluation of some of the technical issues GP relied upon, and

- 3) an analysis of the sensitivity of the 11 mg/kg tPCB footprint size to efforts to correct Aroclors relative to total congeners.

2 PREVIEW OF FINDINGS

The following bullets summarize our findings, and subsequent sections provide the basis for these findings.

- 1) The models developed by GP, for the combined RI and SRI data as well as for the SRI data only, provide a very smooth depiction of generalized features over a broad spatial scale. This approach is not particularly useful for delineating contaminant deposits and evaluating remedial options because the extremes of the data have essentially been filtered out.
- 2) A careful re-evaluation of spatial variation in total Aroclors including both RI and SRI data shows that the ordinary kriging model used to estimate the 11 mg/kg Aroclor footprint was flawed. The semi-variogram was estimated incorrectly and key assumptions necessary for ordinary kriging models are not satisfied in either the SRI data alone or the combined RI and SRI data.
- 3) The natural neighbor model presented at a recent working group meeting failed to correctly represent clear long-flow anisotropy represented in the ordinary kriging models and was also inappropriate for the SRI or SRI and RI combined data. The natural neighbor model presented by GP is very generalized in contrast to the more resolved mapping we present.
- 4) By re-fitting an anisotropic natural neighbor model, we found that the natural neighbor model fit the Aroclor data as well as GP's ordinary kriging model for the SRI data.
- 5) Unlike GP's ordinary kriging model, the anisotropic natural neighbor interpolation presented a more coherent interpretation of the combined RI and SRI data.
 - a. The natural neighbor interpolation reproduced both data sets, accurately identifying highs and lows in the spatial distributions of sample data.
 - b. Despite the Aroclors in the RI data being generally higher than those in the SRI, the ordinary kriging model based on the combined data estimated a smaller lateral extent of 11 mg/kg footprint than the ordinary kriging model based solely on the SRI data. This indicates a flaw in the application of ordinary kriging to the combined data.
 - c. In contrast, when the anisotropic natural neighbor interpolation was applied to the combined data the extent of contamination increased as one would logically expect when combining higher concentration RI samples with the SRI data.
- 6) Because the ordinary kriging model as developed by GP fails to follow basic order relationships linking higher or lower concentrations to larger or smaller remedial footprints, the model should not be relied upon for comparing remedial alternatives in the FS.
- 7) The natural neighbor interpolation and ordinary kriging models had similar cross validation statistics for the combined RI and SRI data, yet the area exceeding 11 mg/kg estimated by natural neighbor interpolation was more than double (116 Acres as compared to 45.6 Acres) when estimated using the ordinary kriging model.
- 8) Laboratory split sample evaluations suggest Aroclor analyses understate total PCBs.

- a. Given the large investment in the SRI data, a factor calibrating Aroclors to total PCBs defined by congeners may be an appropriate step.
- b. We investigated potential inflation of the remedial footprint as a function of correction factors ranging from 1.25 to 2.0 and found that the 11 mg/kg extent is sensitive to even a 25% correction. Footprint size increased more on a percentage basis than the percentage correction of the data.

3 REVIEW OF TECHNICAL WORKGROUP INFORMATION

Until recent work group meetings, analyses provided to the agencies have been based on ordinary kriging models supported by only the SRI data. The models were developed in along-river and across-river coordinates with the major axis of anisotropy oriented along river flow. At the suggestion of the MDEQ, a model was recently presented by GP that included both RI and SRI data, using the same statistical modeling methods, but with unrealistic results for the combined data. The area of floodplain soils exceeding 11 mg/kg Aroclors was estimated to be 55 acres based on the SRI data model but declined to 45 acres for the model supported by RI and SRI data combined. This result was surprising because Aroclor concentrations in the RI data generally exceeded nearby values in the SRI data. GP has concluded that because the combined data do not fit the model developed for the SRI data, the RI data should be excluded. The MDEQ has pointed out that the combined data and potentially the SRI data alone violate key statistical assumptions underlying the ordinary kriging model. The MDEQ has concluded that alternative valid modeling methods are needed to integrate the RI and SRI data in a sensible way. This section documents our approach to re-evaluating the RI and SRI data to understand how the data should be used for development of the FS.

As another line of evidence, we re-estimated the footprints using the natural neighbor method. The natural neighbor method is a technique with essentially no assumptions about the concentration data, other than the interpolated surface match the data at the sampling locations.

We interpolated the Aroclor data using the natural neighbor method in the straightened coordinate system and assuming a 4 to 1 ratio of anisotropy. The natural neighbor method has no specific parameter representing directional anisotropy, so we scaled the long flow coordinates by a factor of 4 so that samples in the long flow direction were interpreted as being a factor of 4 closer together than those in the cross flow direction. We fit two models: the first based on the SRI data (GP, 2014) alone and the second based on the SRI and RI data combined.

Through the technical work group process, GP has recommended that the FS be supported exclusively with the SRI data (ignoring the RI data) or, if the SRI data were to be included, that the SRI data be weighted in some fashion. GP supports this position with four primary analyses.

- 1) General preference for the intuitive appeal of the ordinary kriging model applied to the SRI data exclusively.
- 2) Identifying unsatisfying spatial display of the Aroclors when ordinary kriging or natural neighbor interpolation is applied to the combined data.
- 3) Lower cross validation errors for ordinary kriging model based on SRI data only, as opposed to combined SRI and RI data.

- 4) Semivariogram analyses concluding that, for the combined data, the semivariograms are virtually pure noise—strongly reducing spatial resolution in the mapped values.

The remainder of this section provides a re-evaluation of these points.

3.1 INTUITIVE APPEAL OF SRI ONLY ORDINARY KRIGING MODEL

Maps based on ordinary kriging models with SRI data only provide what appear to be reasonably intuitive spatial patterns in the Aroclor distribution. These patterns include deposits elongated in the direction of flow (as would be expected in a riverine setting) and areas with generally high and low concentrations in the map generally following corresponding patterns in the data (Figure 1, top panel). Conversely, the mapped values developed by GP based on the combined SRI and RI data (Figure 1, bottom panel) are less intuitive with mapped values failing to capture the extremes in the data, even for the SRI data locations. This lack of model to data match and the change in the footprint size (54.6 Acres with SRI data only and 45.6 acres for the combined data; Table 1) is counter intuitive at best and likely fundamentally flawed. GP concluded from this information that there were problems with the combined data because they did not fit the model well. This is the primary basis supporting GP's position to exclude (or possibly weight) the RI data from further consideration. However, it is not clear that the effort to develop a coherent model integrating the two data sets was adequate to draw such important conclusions—potentially leading to the exclusion of hundreds of samples which were also collected by GP and considered completely valid when used to support earlier analyses and Site documents submitted to USEPA for their approval. The USEPA approved terrestrial ecological and human health risk assessments were based on these data. In the next section the underpinnings of the ordinary kriging model are reevaluated to understand their veracity.

Table 1. Mean absolute cross validation errors for ordinary kriging model based on SRI data exclusively and RI+SRI data combined.

Data Subset	Ordinary Kriging Model	
	Mean Absolute Error (mg/kg)	Aroclor >11 mg/kg (Acres)
SRI-Only	2.9	54.6
SRI+RI	4.7	45.6

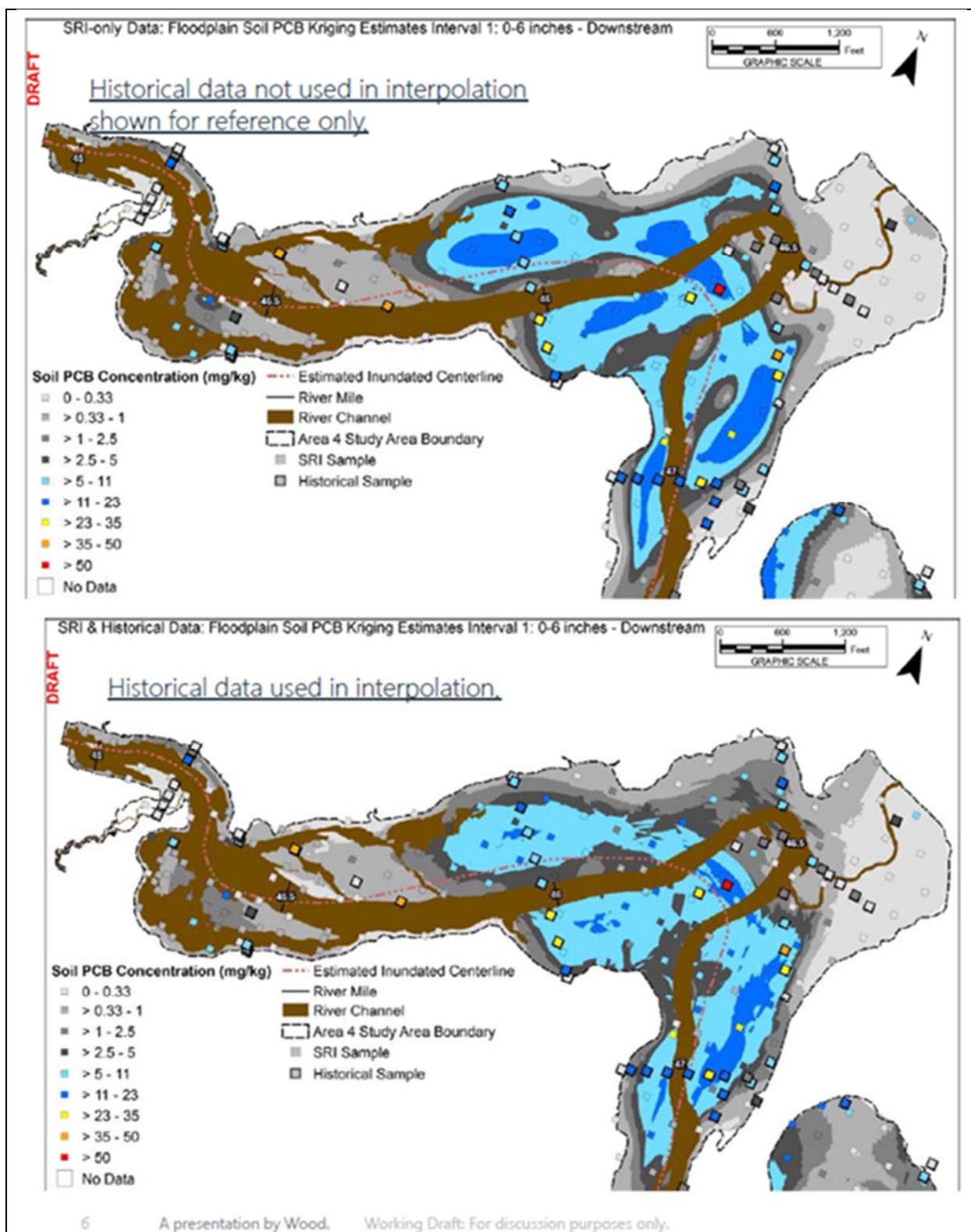


Figure 1. Ordinary kriging model for SRI data (top panel) and RI+SRI combined data (bottom panel) reproduced for technical discussion purposes.

3.2 SEMIVARIOGRAM ANALYSIS

The nature of an ordinary kriging model is primarily determined by parameters derived from analysis of a statistical measure of spatial dissimilarity called the semivariogram (Isaaks and Srivastava, 1989). An example semivariogram is depicted in Figure 2 to illustrate the idealized form and the key pieces of information provided by a semivariogram analysis. The green boxes on the plot are the sample variogram estimates calculated as one half the average of squared differences in concentration for pairs of samples separated by a similar distance. If contaminant concentrations are spatially correlated, points closer together are generally more similar (i.e. lesser semivariogram values) and pairs of points separated by greater distances are generally more dissimilar (i.e. greater semivariogram values). The red line is a fitted semivariogram model and the parameters controlling the shape and position of the model are the three key parameters that serve as inputs to kriging computer programs. These are the Nugget Effect, Sill and the Range of Influence. The semivariogram depicted in Figure 2 has a nugget effect of 0.3 (the vertical axis intercept), a sill of 0.8 (the horizontal asymptote), and a range of influence of about 0.5 distance units (the distance at which the sill height is reached).

The ratio of the nugget effect to the sill height determines how strongly the data influence mapped values. In the extreme situation when the nugget effect is equal to the sill, the interpolated surface is effectively independent of the sample data and all points exert equal influence irrespective of proximity to sample locations. Consequently, the mapped surface is represented by a single value—the sample mean. In the other extreme, when the nugget effect is zero, the mapped kriging surface is forced to exactly match the sample data and the surface generally exhibits much stronger spatial gradients with highs and lows closely mirroring the sample data. Between these extremes, interpolated surfaces can be very smooth when the nugget is close to the sill, or more variable when the nugget is closer to zero. Effectively the ordinary kriging model is a type of weighted least squares fit to the sample data with the leverage each sample exerts on the surface controlled by the nugget to sill ratio and the proximity of mapped locations to sample data.

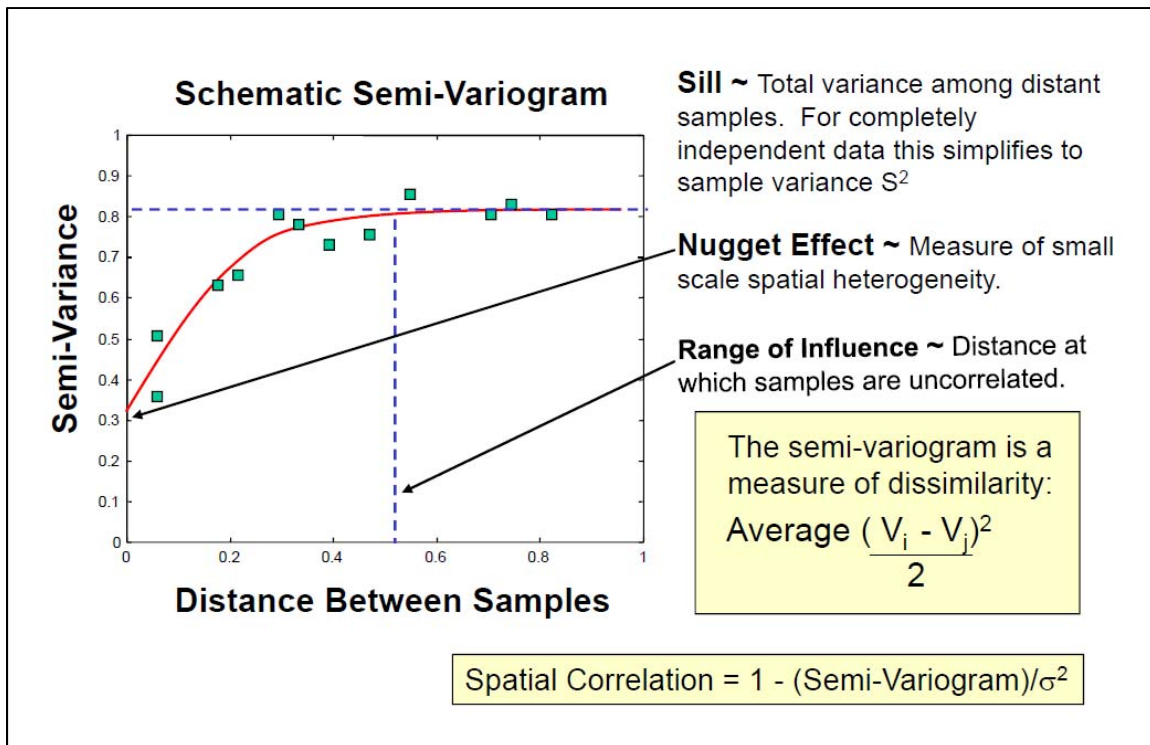


Figure 2. Example semi-variogram estimates and fitted model defining key parameters.

Another key aspect of the semivariogram analysis is estimating how spatial variation may change with direction (anisotropy). Based on broad experience working with contaminated sediment data, it can be expected that samples would be more similar in the upstream to downstream direction (along flow) than in the cross flow direction. This has been noted by GP in their analyses at the Kalamazoo River. To estimate directional semivariograms, paired sample differences are averaged not only within distance groups, but also within sectors of the directional compass. Ideally the long flow semivariogram would be estimated exclusively with pairs of points that are exactly upstream and downstream of each other; however, few pairs ever align so tightly. Therefore, semivariogram estimates are developed through a compromise balancing the narrowest angular tolerance between pairs and an adequate sample size to estimate the semivariogram reliably. A schematic illustrating how pairs are grouped by distance and direction for estimating a directional semivariogram is shown in Figure 3

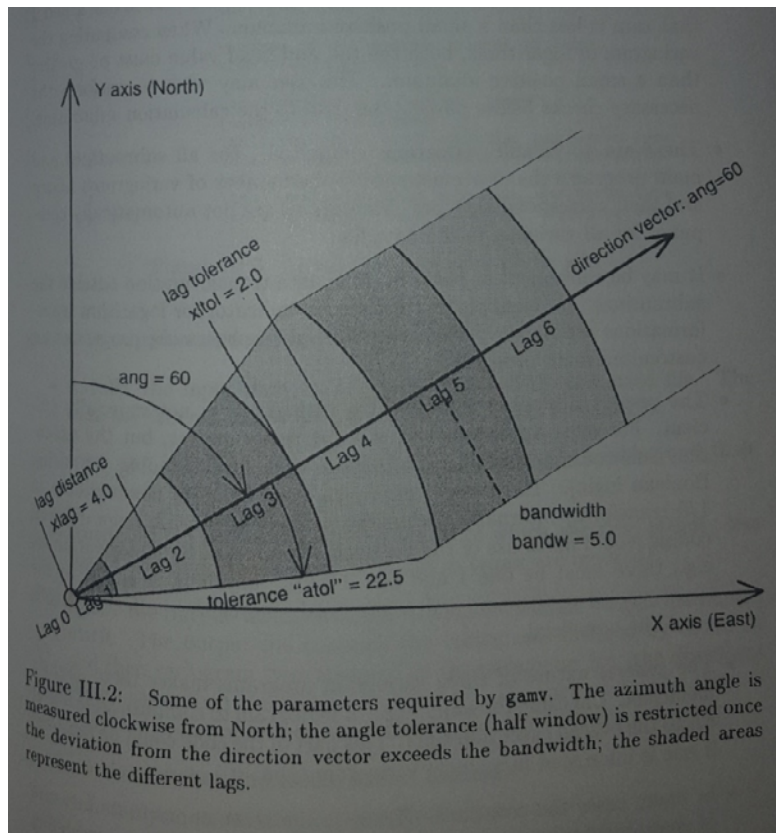
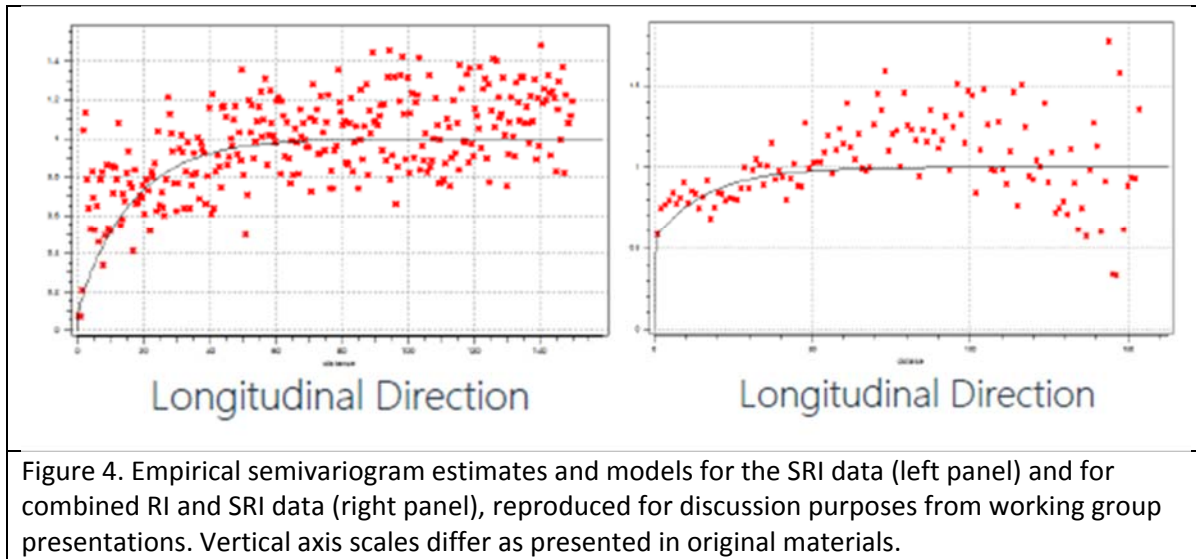


Figure 3. Sample semivariogram grouping options for estimating a directional semivariogram, reproduced from Deutsch and Journel (1998).

Semivariograms presented to the working group showed that for the SRI data there was essentially no nugget effect (Figure 4, left panel) where the fitted model went through the origin. However, for the model based on SRI and RI data combined, the nugget effect was more than 50% of the sill. This large nugget to sill ratio is the reason why the kriging model fails to strongly reproduce the data, which would be a correct interpretation of the data if this long-flow semivariogram is an accurate estimate of the semivariogram model.

Ordinary kriging is a parametric statistical procedure that requires the assumption that the semivariogram is the same for all ranges of concentrations. This means that spatial correlation for values in the extreme upper and lower tails should be no different than for values near the mean of the contaminant distribution. It is generally understood that this assumption (second order stationarity) is unlikely for contaminant distributions because extreme values usually occur much less predictably and are less reproducible than values that are close to the mean concentration. Violation of this assumption can cause estimates of the nugget effect to be biased high because the “average” small scale variation is dominated by a small number of samples in close proximity that also differ strongly. Kriging maps based on such semivariograms are likely to be overly generalized and lacking in spatial resolution, particularly in proximity to the spatial gradients in concentration, precisely where accuracy may be most important when delineating contaminant deposits.



To understand this potential, the MDEQ suggested a nonparametric approach which relaxes the assumption of a single semivariogram and specifically models spatial continuity separately for ranges of contaminant concentrations. GP developed such semivariograms, but they appeared overly general and so the MDEQ has reevaluated these estimates. Through careful estimation of the nugget effect (separately from the large scale variation) and by tightly restricting directional grouping constraints, semivariogram models were developed for 1-, 5-, 11-, 20-, and 25 mg/kg thresholds. Together these semivariograms present a substantively different interpretation of how surface Aroclors vary spatially over the site.

Figure 5 provides an illustration of the differences in semivariogram models. Each is estimated from the combined RI and SRI data, but with apparently differing grouping and averaging approaches. Our estimated model shows that the nugget is just 14% of the sill, as compared to 50% estimated by GP. This indicates that small scale variation is much less than that estimated by GP and that interpolated surfaces should much more closely match sample data near sampled locations. This is in direct contrast to the behavior of maps developed by GP based on semivariograms that appear to overstate the nugget effect. The models developed by GP for the SRI data and for the combined RI and SRI data both provide a very smooth depiction of generalized features over a broad spatial scale. This is not particularly useful for delineating contaminant deposits and evaluating remedial options because the extremes of the data have essentially been filtered out.

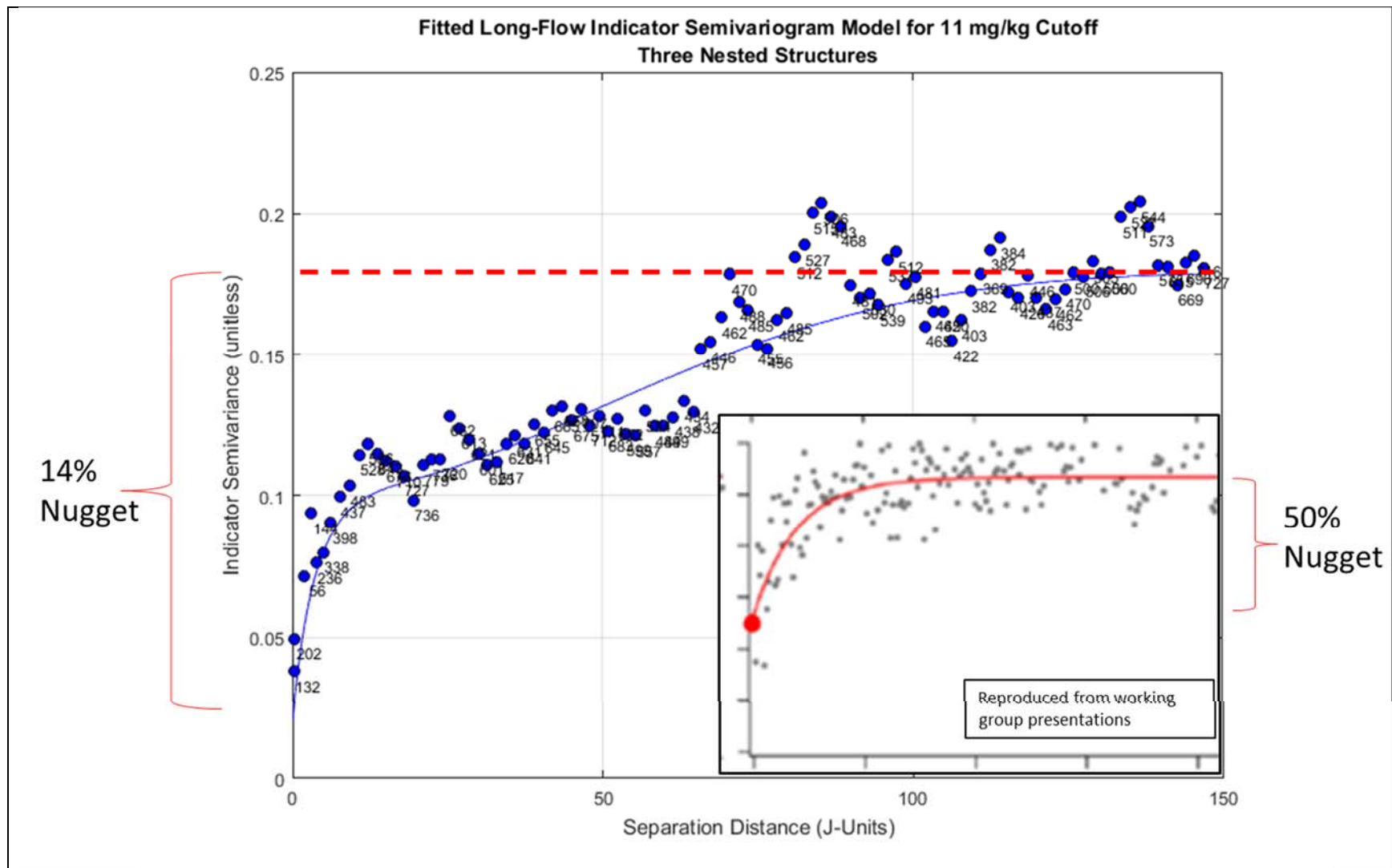


Figure 5. Semivariogram estimates and fitted model for the long flow direction with a range of influence of approximately 125 units and a 14% nugget effect. Inset provided for comparison to GP model presented at working group meetings indicating 50% nugget.

3.2.1 Ratio of Anisotropy

GP presented semivariograms which represent long-flow and cross-flow ranges of influence, which understate the degree to which Aroclor deposits should be elongated along flow for both their ordinary kriging and natural neighbor models. The ratio of the long-flow to cross-flow ranges of influence determine the elongation of Aroclor deposits shown on maps. This is important because judging the degree to which RI and SRI data are compatible for mapping requires that interpolated values properly weight pairs of points along flow in comparison to how pairs of points in the cross-flow direction are weighted.

The data are analyzed in what GP terms (I,J) space with the J-coordinate representing the cross flow direction. We noted that the J coordinate ranges from a minimum of -20 units to a maximum of +20 units, so the longest distances that can be compared in the cross flow direction are 40 units apart (Figure 6). However, in GP's presentation materials the longest distances shown are approximately 80 units. Pairs of points that are 80 units apart must be oriented at more than 45 degrees off the cross flow direction. Effectively pairs of points far from the cross flow direction are averaged together which biases the semivariogram estimates (Figure 6). Because concentrations are more strongly correlated in the long-flow than cross-flow direction, this broad directional pooling of the empirical semivariogram estimates causes the range of influence to be overstated in the cross-flow direction and understated in the long-flow direction. Taken together, the ratio of anisotropy (directional elongation of deposits) is understated and decreases accuracy of the maps.

When we narrowly grouped along the cross flow direction, we found much shorter ranges of influence, suggesting a ratio of anisotropy on the order of 4 to 1 (Figure 7). This larger ratio of anisotropy is consistent with the physical setting. In a riverine setting, the range of influence usually reaches its maximum (the sill) at approximately half the distance across the river or, in this case, floodplain. Concentrations on either edge of the floodplain are generally more similar (lower semi-variogram) than pairs of concentrations where one is on the edge of the floodplain and the other is closer to the river banks. Our cross-flow semivariogram exhibits this typical behavior with the strongest dissimilarity at about half the distance across the J dimension (Figure 7, right panel) and sample pairs separated by over 30 units, essentially opposite edges of the floodplain being as similar as points less than one unit apart. The cross flow semivariograms presented by GP are biased toward greater spatial continuity and as a result total Aroclor deposits are depicted as being too discontinuous along flow and too well connected across flow. Our natural neighbor interpolation provides a contrasting and, we believe, more accurate depiction with 4 to 1 ratio of anisotropy. Our position is discussed further in Section 4 below.

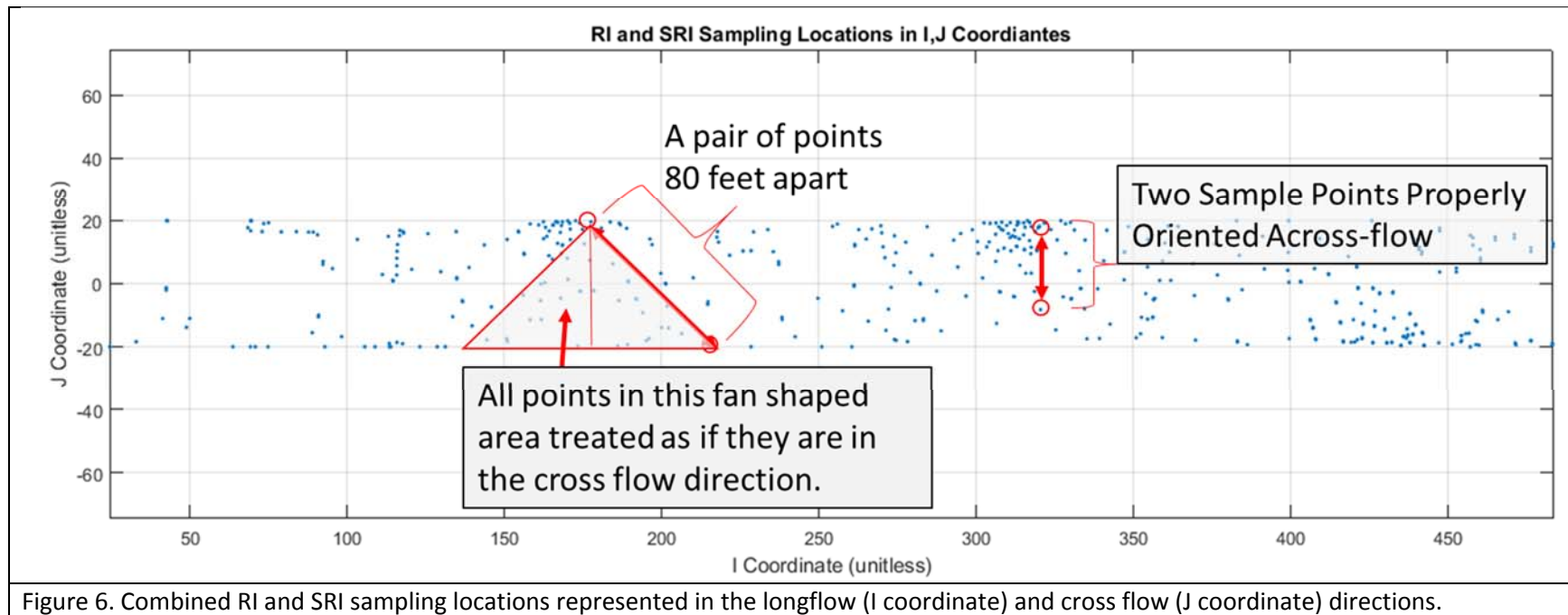
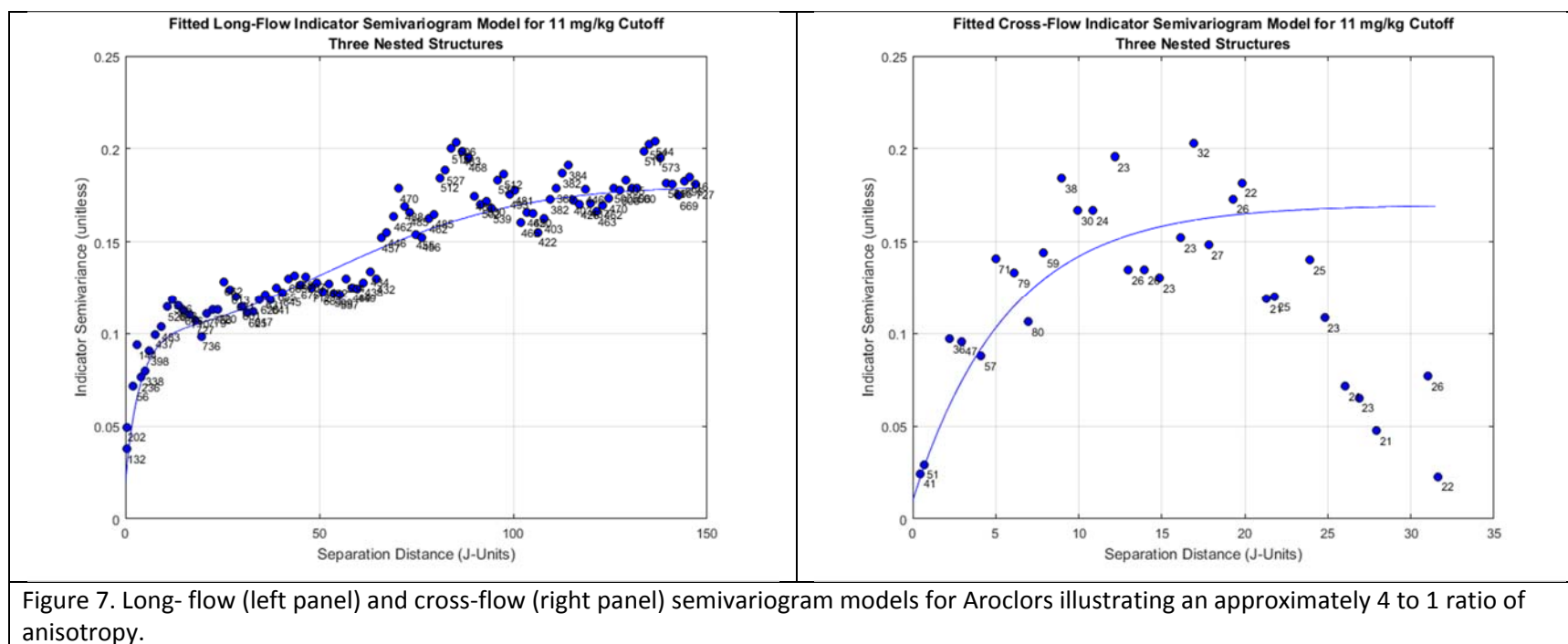


Figure 6. Combined RI and SRI sampling locations represented in the longflow (I coordinate) and cross flow (J coordinate) directions.



3.2.2 Which Model to Trust

Referring back to Figure 3 is a good place to start for understanding which semivariogram models more accurately reflect the actual spatial variation of Aroclors in surface floodplain soils. When estimating the directional semivariogram, any averaging of sample pairs that are not strictly along the principal direction of strongest spatial correlation has the effect of overstating the nugget and understating the range of influence. Generally speaking, all semivariogram analyses based on data that are not perfectly aligned along directional axes are biased estimates. So, while the semivariogram we estimated is based on carefully controlled grouping of samples along the principal axis of correlation, it nonetheless should be considered to understate the sill and overstate the nugget, but apparently less so than the semivariogram presented by GP. Short of running low on numbers of pairs for averaging, it is impossible to understate the nugget or overstate the range of influence. It appears that the semivariograms presented by GP failed to carefully constrain the directional grouping of samples, significantly overstating the nugget effect and to a lesser degree understating the range of influence.

3.2.3 Role of the Semivariogram in Kriging

The semivariograms we estimated had much lower nugget effect than those presented previously by GP. This indicates that mapped concentrations should exhibit much stronger spatial gradients and that mapped values should very nearly match measured concentrations at the sampled locations. This is in stark contrast to the ordinary kriging model which fails to reflect the obvious spatial gradients in sample measurements.

For ordinary kriging, the semivariogram model determines how closely the interpolated surface matches the sample data and how abruptly the surface changes in areas where there may be steep concentration gradients. For low nugget and long ranges of influence, interpolated surfaces vary smoothly between sampled locations exactly matching the sample data. When the nugget is a large percentage of the sill, the model interprets data as if it is measured with substantial measurement error and the interpolated surface varies much less and the average of sample values within a local area dominates the mapped concentrations, even in close proximity to measured locations. The semivariograms estimated by GP for the combined RI and SRI data have very high percentage nugget effect and the kriged maps are nearly constant throughout the area because mapped values predominantly reflect large area averages rather than local patterns. Our semivariograms that we estimated by carefully studying the nugget effect and directional pairing of sample data directly contradict GP's interpretation of the Aroclor distribution.

3.2.4 Practical Implications

The modeled surfaces provided to the working groups fail to accurately reflect local spatial concentration gradients that are apparent in the sample data. Because mapped surfaces are very generalized they are likely to understate the range of exposures that small home range receptors are likely to experience. This will in turn tend to result in a low bias in the estimated remedial footprint and inaccurate evaluation of remedial alternatives.

Mapping methods that are not reliant on a second order stationarity assumption such as natural neighbor interpolation, or indicator kriging with properly estimated indicator semivariograms would more accurately reflect the statistical distributions and spatial variation in the sample data. In

subsequent sections we investigate maps based on natural neighbor interpolation to further understand the practical importance of the mismatch between the ordinary kriging model and the more variable distribution of Aroclors indicated by our semivariogram analysis.

4 ALTERNATIVE MAPPING

The ordinary kriging model based on the semivariograms developed by GP do not provide a satisfying map of Aroclors in Area 4. Our analyses suggest that an alternative approach is needed to more accurately reflect the spatial distribution of the combined RI and SRI data. GP has proposed eliminating or weighting hundreds of RI data, most of which was collected by GP. Absent some technical rationale for removing or weighting the data, this is an untenable solution for the regulatory agencies to defend. In this section we present an alternative mapping method based on natural neighbor interpolation which has no parametric assumptions of stationarity. Our approach also more closely matches the sample data, which is more consistent with the low nugget effect we estimated with our semivariogram analysis. Ideally we would develop an indicator kriging model for this purpose, but in the interest of time the simpler natural neighbor model was evaluated. Further evaluations may include development of a conditional simulation model based on indicator kriging so that the uncertainty in key quantities of the feasibility study can be evaluated rigorously.

4.1 MODEL FIT (CROSS VALIDATION)

For this analysis, we developed maps based on the SRI data alone (Figure 8) as well as the combined RI and SRI data (Figure 9). We then compared these maps with the quantitatively and qualitatively with ordinary kriging based maps. The quantitative review included evaluation of model fit as well as differences in the size of the area exceeding 11 mg/kg¹.

We tested the model fit by sequentially dropping one sample at a time and using the remainder of the data to predict the sample value at the dropped location, a method known as cross validation (Isaaks and Srivastava, 1989). The difference between actual and predicted value is the residual error and we summarized model performance by calculating mean absolute value of these errors (MAE). We calculated the MAE for each of the two data groups and compared the MAE for each data configuration with corresponding ordinary kriging models fit by GP to the same data configurations. We also calculated the area of floodplain soils exceeding 11 mg/kg Aroclors. We compared these estimates with the corresponding estimates reported by GP based on kriging models.

GP reported mean absolute errors of 2.9 mg/kg and 4.7 mg/kg for ordinary kriging models and the natural neighbor model fit the sample data equivalently with mean absolute errors of 2.9 mg/kg and 5.1 mg/kg respectively for each data configuration (Table 1). This equivalence of model fit indicates that all

¹ The 11 mg/kg threshold was selected because this is the site's risk-based concentration for organisms with 1 to 2 acre home ranges, such as robins and shrews. We recognize that the feasibility study will be based on quantities derived from these maps, but practically speaking the size the remedial footprint based on a moving window analysis will be approximately proportional to the 11 mg/kg footprint.

other things being equal, either interpretation of the data can be considered equally likely. However, the stronger local agreement between data and the natural neighbor interpolation is in closer agreement with the semivariogram analysis described above. The low nugget effect we estimated indicates that the interpolated surface should be more responsive to the sample data, in contrast to the tendency for ordinary kriging model to be drawn to the overall mean of the data. From a weight of evidence perspective, results based on the natural neighbor method are likely to more closely reflect the actual situation in the field because the method does not require the stationarity assumptions of ordinary kriging and the spatial patterns are more like those that can be inferred from our semivariogram analysis.

4.2 AREA EXCEEDING 11 MG/KG AROCLORS

The natural neighbor interpolation and the ordinary kriging models each fit the sample data similarly (as discussed above), with nearly identical cross validation statistics. However, the estimated area exceeding 11 mg/kg was starkly different between the two methods. For the SRI data alone, the ordinary kriging model identified 54.6 acres exceeding 11 mg/kg, whereas the natural neighbor interpolation identified 74 acres exceeding 11 mg/kg, a nearly 50% difference solely based on how the data were modeled (Table 2). When the two methods were applied to the combined RI and SRI data, the ordinary kriging model indicated a smaller area exceeding 11 mg/kg whereas the natural neighbor model logically identified a larger footprint (116 acres) when the higher concentrations found in the RI were included. This amounts to more than a 100% difference between the two estimates when using the combined RI and SRI data, despite both modeling methods producing nearly identical cross validation statistics.

Table 2. Model Fit and Area Estimates for PCB Concentration Exceeding 11 mg/kg Total Aroclors

Data Subset	Kriging Model		Natural Neighbor Model	
	Mean Absolute Error (mg/kg)	Area with PCB > 11 (Acres)	Mean Absolute Error (mg/kg)	Area with PCB > 11 mg/kg (Acres)
SRI-Only	2.9	54.6	2.9	74
SRI+RI	4.7	45.6	5.1	116

Figure 8 and Figure 9 show the resultant natural neighbor interpolation models for SRI data only as well as for combined RI and SRI data. For comparison, the ordinary kriging models presented by GP are also included as insets in each figure. The Natural neighbor interpolation was estimated with the assumption of a 4 to 1 ratio of long flow to cross flow range of influence. It can be seen that the resultant map is intuitively correct in that deposits are elongated with flow and also the areas with elevated concentrations tend to track the locations where higher concentration samples are found. The model is qualitatively similar to the ordinary kriging model with the primary exception being that the ordinary kriging model tends to be more generalized with areas exceeding 11 mg/kg (dark blue on GP's map included as an inset) being less spatially continuous and somewhat smaller in total area. Both models fit the data equally well, but the area estimates are meaningfully different with the natural neighbor model

indicating 50% more area exceeding 11 mg/kg. This difference in area exceeding 11 mg/kg is likely to substantively impact costs estimated in the feasibility study.

When fit to the RI and SRI data combined, it is clear that the ordinary kriging model is inaccurate. The ordinary kriging model fails to reflect the additional data with generally higher concentrations (central tendency and extremes) than the SRI data. Figure 9 shows that the natural neighbor provides a more plausible depiction of the contaminant distribution conditional on all of the available data. It is also notable that addition of the RI data to the SRI data does not cause deposits to change locations, but rather the spatial resolution and spatial gradients are sharpened with their inclusion. While both models exhibited similar cross validation statistics, it is important to note that the derived quantities such as remedial footprints for the natural neighbor model are likely to more accurately represent field conditions because it preserves the short scale continuity which we found in the corrected semivariogram models presented above. In contrast, the ordinary kriging model largely breaks down because it is based on inappropriate semivariogram models and because the data do not satisfy the parametric assumptions of second order stationarity. Based on our analysis, the semivariogram clearly changes with concentration - values below 0.33 mg/kg and greater than 15 mg/kg are much less strongly spatially correlated than areas with concentrations closer to 11 mg/kg. Choice of modeling methods is only one of several issues that may contribute to uncertainty in the remedial footprint. In the next section, sensitivity to analytical bias is evaluated, an issue that could explain apparent differences between RI and SRI data.

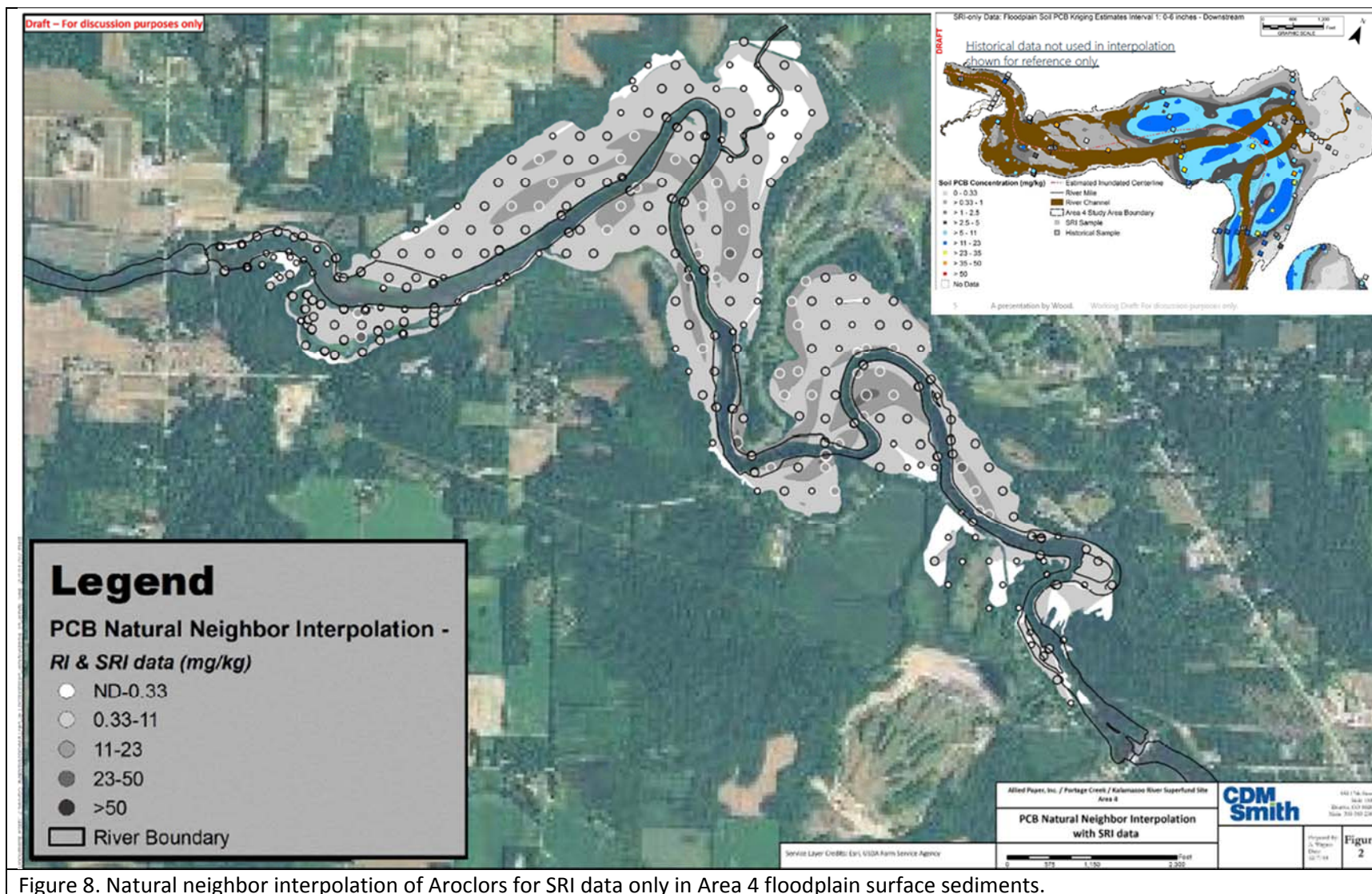


Figure 8. Natural neighbor interpolation of Aroclors for SRI data only in Area 4 floodplain surface sediments.

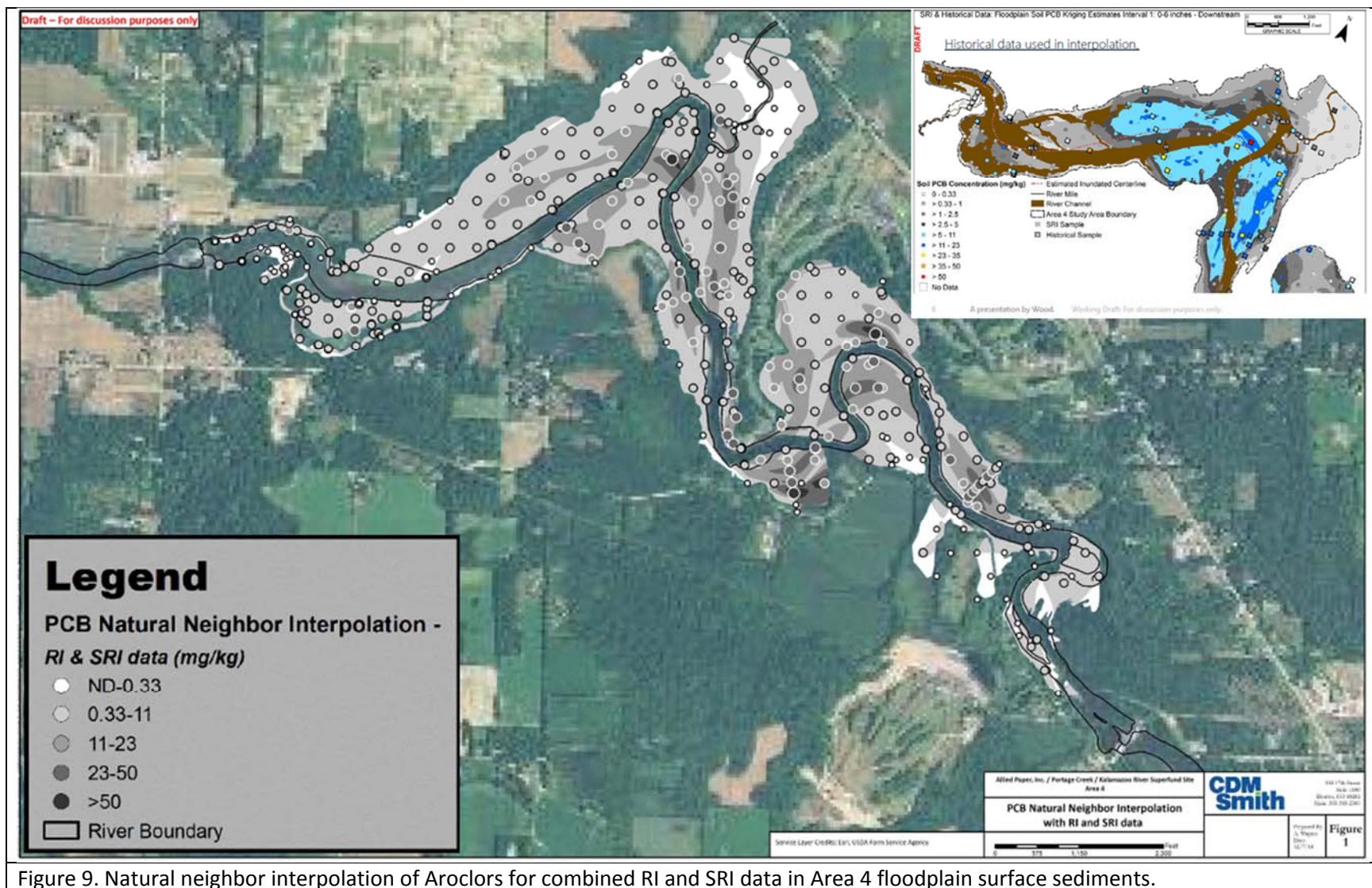


Figure 9. Natural neighbor interpolation of Aroclors for combined RI and SRI data in Area 4 floodplain surface sediments.

5 SENSITIVITY TO AROCLOR CORRECTION FACTORS

Previous analyses comparing RI and SRI data suggest that Aroclor totals from the SRI data may understate total PCBs when compared to more accurate congener analyses. Preliminary analyses comparing laboratory split samples, analysis of CRM, and splits comparing Aroclor totals with congener totals suggest that the SRI Aroclor totals may be biased low. Total PCB congeners may exceed Aroclors reported by GP by a factor of 1.25 to 2.0. None of the split sample comparisons in and of themselves is conclusive, but all indicate that Aroclor analyses tend to understate total PCBs defined as congener totals and that the SRI data are biased lower than the RI data. In this section the natural neighbor method, which was established in the previous section as providing a robust depiction of the spatial distribution of Aroclors for SRI data alone and for combined RI and SRI data, is used to explore the effects of correcting the data for bias relative to total PCB congeners. Because of this additional robustness to integrating apparently disparate data sets and because of the minimal statistical assumptions, we view these comparisons to be the most accurate basis for comparison of scenarios. Ultimately, more rigorous statistical comparisons should be developed with indicator kriging based conditional simulation methods.

If these apparent analytical differences between Aroclors and total PCB congeners are resolved prior to the record of decision (ROD) and remedial design, we believe there are two possible outcomes:

- 1) If the SRI data are found to be accurate, and the RI data are determined to be biased high, the remedial footprint would be developed based on the SRI data only and future samples that are equally accurate. In this case, the SRI based estimate of the remedial footprint size would be on the order of 50 to 75 acres and there would be no need to make any correction factors.
- 2) Alternatively, if the Aroclors are found to understate total PCBs (as compared to total PCB congeners), a relationship would probably be developed to predict total PCB congeners from the spatially extensive SRI data and presumably future design samples. In this case the SRI data would be corrected and new maps would be developed based on corrected SRI data combined with the RI data, which could be argued to not need substantial correction—the RI concentrations are already higher than the SRI concentrations.

Under the second scenario the sensitivity of footprint estimates to correcting SRI data depends on how close existing sample values are to the 11 mg/kg threshold. If a substantial number of SRI samples are near the 11 mg/kg threshold, footprint size could change substantially, even with a modest 25% correction factor.

To understand sensitivity of the remedial action footprint size due to correcting the Aroclor totals for analytical bias, we applied correction factors 1.25, 1.5 and 2.0 to the SRI Aroclor data and estimated the remedial footprint size based on the anisotropic natural neighbor interpolation for corrected SRI data combined with RI data. The estimated footprint sizes are summarized in Table 3 where the footprint size for a 25% correction (1.25 multiplier) increases from an estimated 116 acres to 148 acres, a 27% increase over our estimate based on natural neighbor, and a 171% increase relative to the 54.6 acres estimated by GP using ordinary kriging based on the SRI data only. Increasing correction factors result in substantially larger footprint sizes ranging up to 206 acres for a correction factor of 2.0. We currently

have no reason to expect the Aroclor data to be this heavily biased relative to total PCB congeners, but a 25% or 50% error is not out of the question. There is also no real basis to argue that the Aroclor totals reported in the RI are not also biased low, although the data to determine this may not be available.

Table 3. Sensitivity of remedial footprint estimates to correction of supplemental RI data.			
Correction Factor Applied to SRI Data	Area With Surface Total Aroclors Exceeding 11 mg/kg (Acres)	Increase Relative to Natural Neighbor ¹	Increase Relative to Ordinary Kriging ²
1	116	NA	112%
1.25	148	27%	171%
1.5	171	47%	213%
2	206	77%	277%

Notes:

- 1) Area for the natural neighbor interpolation was based on the combined RI and SRI data.
- 2) Area with surface total Aroclors exceeding 11 mg/kg based on the ordinary kriging model was 54.6 acres based solely on the uncorrected SRI data.

Figure 10. Natural neighbor interpolation of total Aroclors for combined RI and SRI data with a 25% correction (Factor of 1.25) applied to the SRI data in Area 4 floodplain surface sediments. depicts the natural neighbor interpolation of the RI and SRI Aroclor totals with the 25% correction applied to the SRI data. The resultant map is qualitatively similar to the uncorrected map shown in Figure 9 with the highs and lows in the same general places, but with more spatial continuity of Aroclor deposits in the corrected map, greater lateral extent of the deposits, and higher peak values with some modest areas mapped as exceeding 50 mg/kg and substantial areas with concentrations between 23 mg/kg and 50 mg/kg. Although we have not conducted the moving home range analysis thus far, this map suggests that the number of home ranges with average concentrations exceeding 11 mg/kg is likely to differ substantially from the ordinary kriging map depicted in the Figure 8 inset above. Figure 11 and Figure 12 present maps of Aroclors based on the natural neighbor interpolation with combined RI and corrected SRI data for factors of 1.5 and 2.0, respectively. As indicated in Table 3. Sensitivity of remedial footprint estimates to correction of supplemental RI data. the areas exceeding 11 mg/kg are increasingly larger and more spatially contiguous as the degree of correction increases.

Taken as a whole, these results indicate that the area exceeding 11 mg/kg is meaningfully sensitive to the application of a correction factor to the SRI data and a correction of the RI data would exacerbate this sensitivity. GP currently estimates the area exceeding 11 mg/kg to be approximately 55 acres. By comparison we have found that integrating RI and SRI data using models consistent with the statistical properties of the sample data results in meaningfully larger footprints with a plausible range of 116 to 148 acres which is a factor of 2 to 3 greater than the estimate currently being put forth by GP.

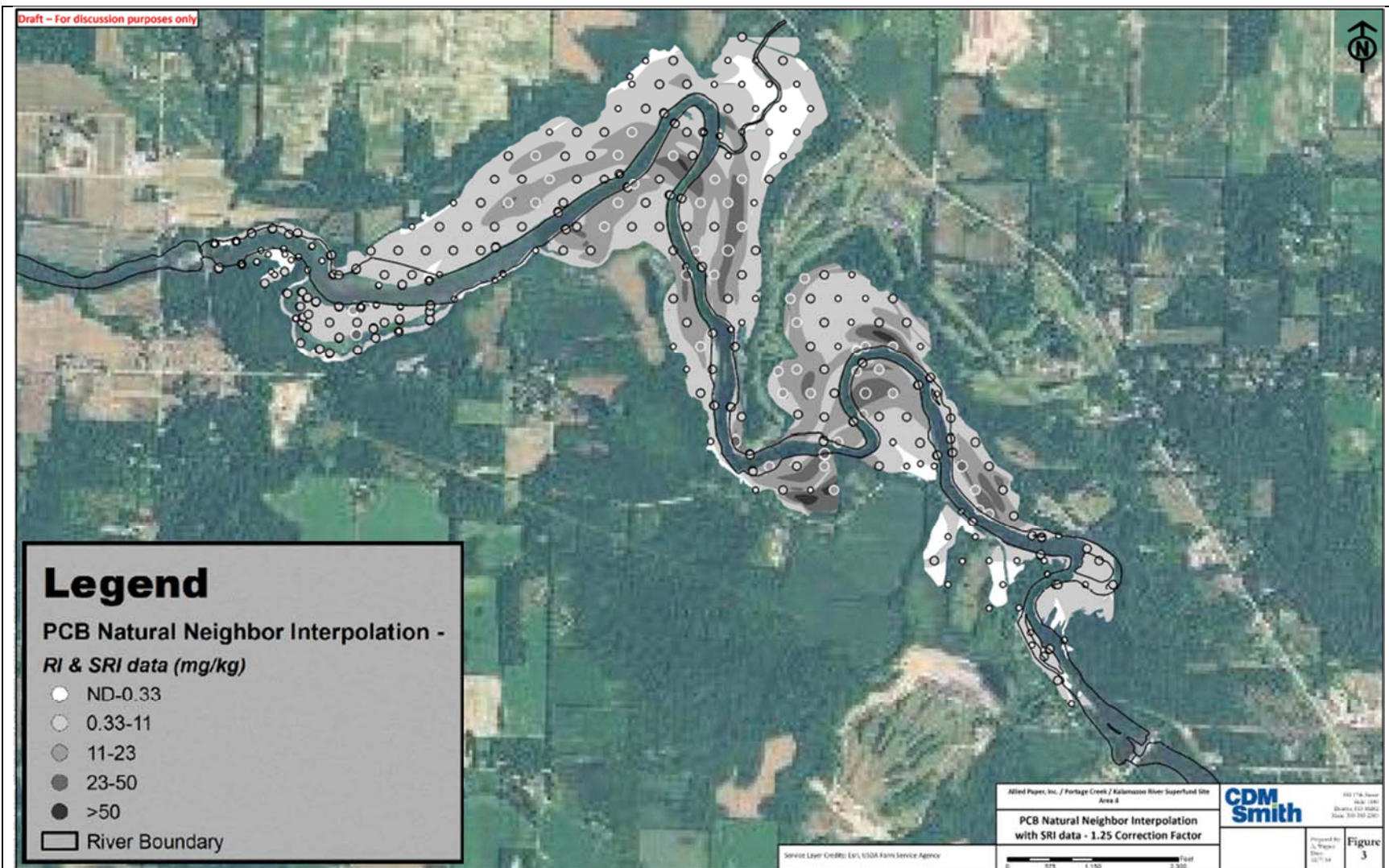
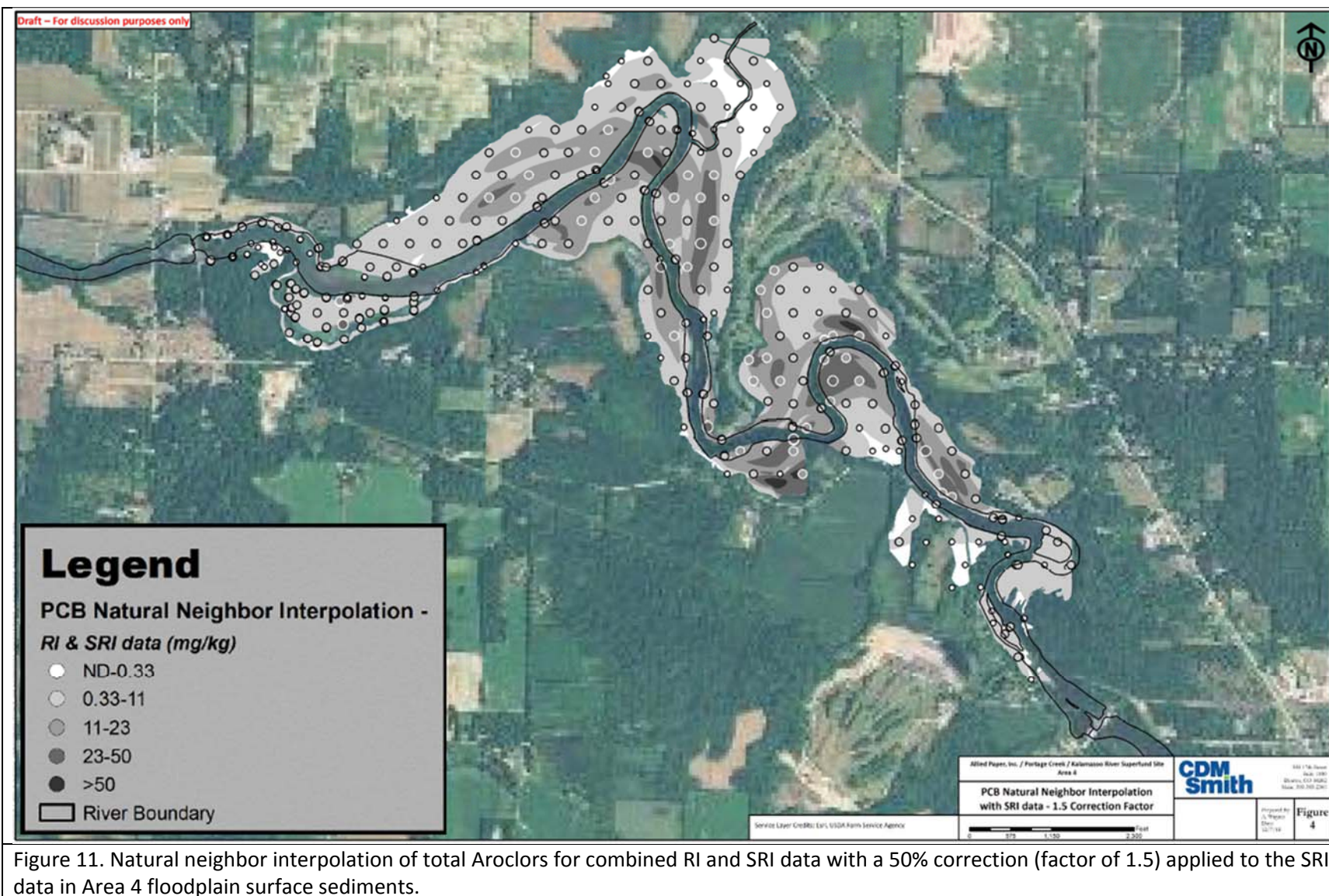


Figure 10. Natural neighbor interpolation of total Aroclors for combined RI and SRI data with a 25% correction (Factor of 1.25) applied to the SRI data in Area 4 floodplain surface sediments.



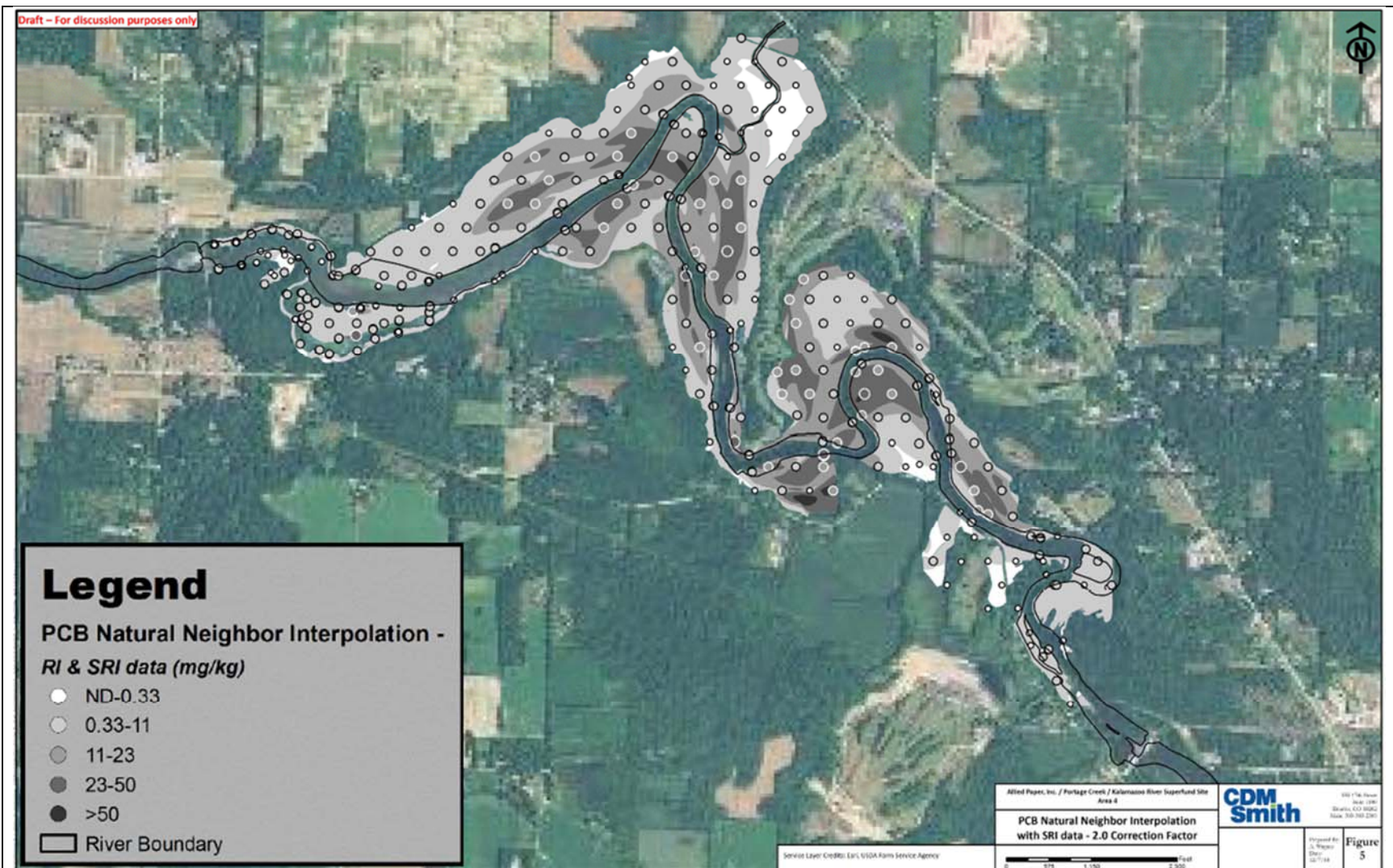


Figure 12. Natural neighbor interpolation of total Aroclors for combined RI and SRI data with a 100% correction (Factor of 2) applied to the SRI data in Area 4 floodplain surface sediments.

6 DISCUSSION AND CONCLUSIONS

In our analysis, we found that ordinary kriging models and natural neighbor interpolation models fit the data equally well yet resulted in substantively different estimates of the area exceeding 11 mg/kg. For the combined RI and SRI data, the natural neighbor was robust in that it behaved in a satisfying way qualitatively, and quantitatively provided a stable method to evaluate some scenarios that may be encountered when apparent biases between data sources is resolved.

Previous analyses suggest that the SRI data may be biased low relative to other laboratory Aroclor analyses, CRMs, and total PCB congeners. When we evaluated the sensitivity of the remedial footprint size to correction of these possible biases, all evaluations (including use of the natural neighbor model rather than ordinary kriging or applying correction factors to the analytical data) found that the 11 mg/kg footprint was meaningfully larger than the 55 acres estimated by GP. We believe based on these analyses that the 55-acre estimate is likely the lower bound of what may be encountered when new data are collected for remedial design and apparent analytical biases are resolved. At a minimum, we see the 116 acres we estimated from the combined RI and SRI data as a reasonable estimate but also believe that this may not be an upper bound. Any correction of one or both data sets to match higher total PCB congener values would increase the footprint as well. We see the 147 acre estimate that reflects a 25% bias correction of the SRI data as a reasonable possibility for an upper bound.

When two models fit sample data equally but lead to meaningfully different management implications, the data and supporting information are likely too uncertain to distinguish between two equally supported management actions. In the statistical literature this is termed model uncertainty—the correct model formulation is unknown and the data do not provide adequate information to identify it. Based on our analysis, we believe this model uncertainty far exceeds CERCLA standards that remedial costs which should be estimated to within -30% or +50% relative errors. With potential for 100% or more deviation between high and low estimates of the area exceeding 11 mg/kg we do not believe these standards can be met without resolution of the discrepancies between RI and SRI data. To date, there has been no information developed to suggest that the RI data, collected by GP consultants, are inaccurate. It is the MDEQ position that either issues related to data usage need to be resolved before the feasibility study is conducted for Area 4, or that the feasibility study explicitly incorporate these uncertainties in a rigorous error analysis including scenarios reflecting the potentially corrected data set(s). The FS and ROD also need to include language spelling out procedures that will be followed to resolve these discrepancies.

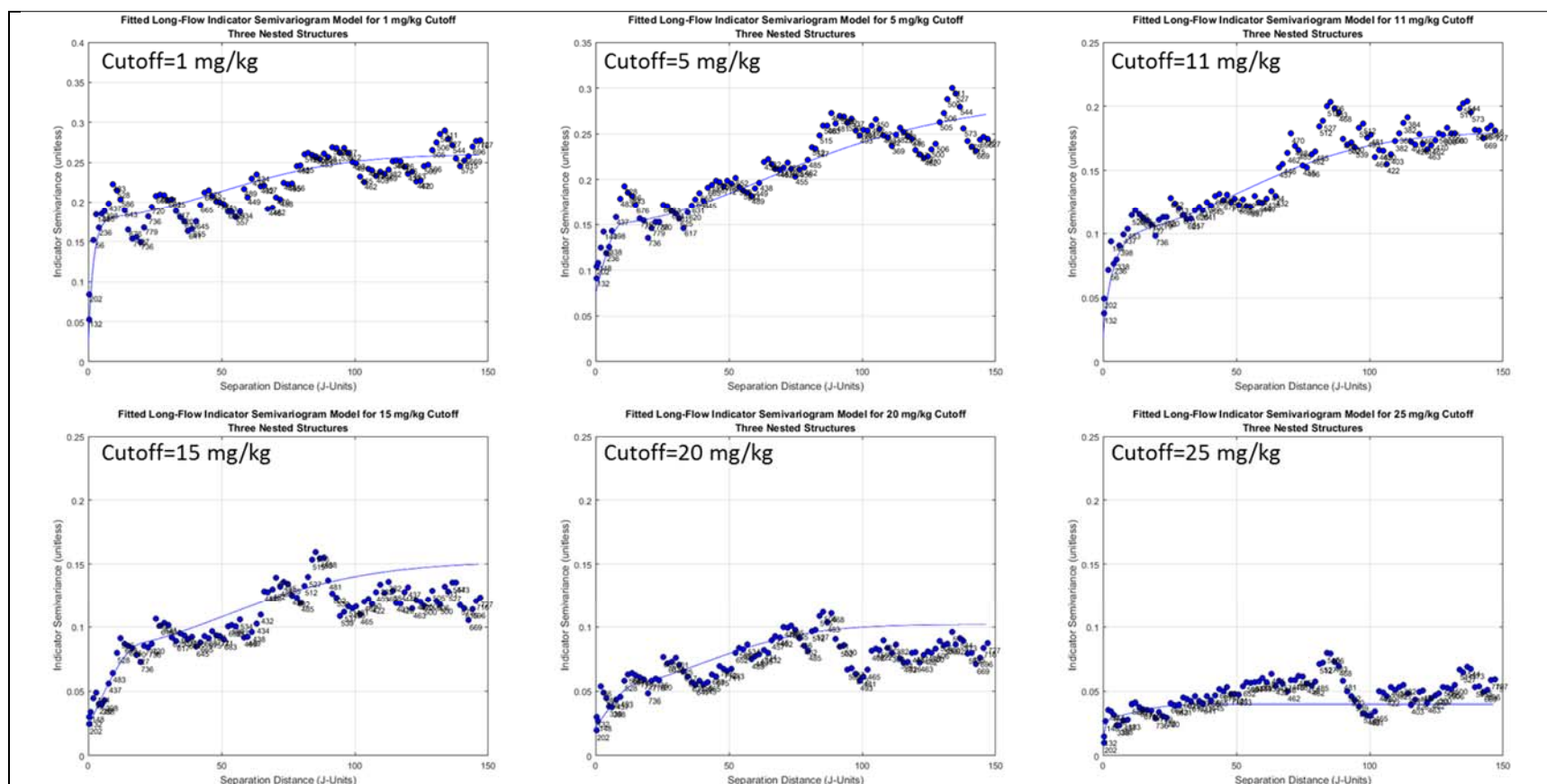


Figure 13. Empirical long-flow semivariogram estimates and fitted models for total Aroclors based on RI and SRI data combined. Nugget effect was estimated considering pairs of samples within 0.28 units and 0.44 units irrespective of direction and remaining pairs of points were constrained to be within an angular tolerance of 12 degrees (atoll parameter in GSLIB) and a directional band width (bandh parameter in GSLIB) of 2 units.

7 REFERENCES

- CDM Smith. 2018a. Working Draft Summary of Area 1 Split Sample Evaluation. Memorandum. October 15.
- CDM Smith. 2018b. Summary of Allied Paper, Inc./Portage Creek/Kalamazoo River Superfund Site, Operable Unit 5, Area 1 Polychlorinated Biphenyl Congener and Aroclor Split Soil Samples. Memorandum. November 28.
- EPA. (1999). Method 1668, Revision A: Chlorinate Biphenyl Congeners in Water, Soil, Sediment, and Tissue by HRGC/HRMS. Washington, DC: United States Environmental Protection Agency.
- Isaaks, E., and R.M. Srivastava. 1989. *Introduction to Geostatistics*. Oxford University Press. New York.
- Kern Statistical Services, Inc. 2001. Allied Paper Inc./Portage Creek/Kalamazoo River Superfund Site. USEPA Removal Assessment: Stage One Data Evaluation, Plainwell/Otsego City Impoundments. Unpublished Technical Research Report.